



# Discovering Putative Prion-Like Proteins in *Plasmodium falciparum*: A Computational and Experimental Analysis

Irantzu Pallarès<sup>1,2†</sup>, Natalia S. de Groot<sup>3,4†</sup>, Valentín Iglesias<sup>1,2†</sup>, Ricardo Sant'Anna<sup>1,2</sup>, Arnau Biosca<sup>5,6,7</sup>, Xavier Fernàndez-Busquets<sup>5,6,7</sup> and Salvador Ventura<sup>1,2\*</sup>

## OPEN ACCESS

### Edited by:

Rustam Aminov,  
University of Aberdeen,  
United Kingdom

### Reviewed by:

Nyssa Drinkwater,  
Monash University, Australia  
Thomas Dandekar,  
Universität Würzburg, Germany

### \*Correspondence:

Salvador Ventura  
salvador.ventura@uab.es

† These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Infectious Diseases,  
a section of the journal  
Frontiers in Microbiology

**Received:** 23 April 2018

**Accepted:** 11 July 2018

**Published:** 07 August 2018

### Citation:

Pallarès I, de Groot NS, Iglesias V,  
Sant'Anna R, Biosca A,  
Fernàndez-Busquets X and  
Ventura S (2018) Discovering Putative  
Prion-Like Proteins in *Plasmodium*  
*falciparum*: A Computational  
and Experimental Analysis.  
Front. Microbiol. 9:1737.  
doi: 10.3389/fmicb.2018.01737

<sup>1</sup> Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain, <sup>2</sup> Departament de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, Barcelona, Spain, <sup>3</sup> Centre for Genomic Regulation, The Barcelona Institute of Science and Technology, Barcelona, Spain, <sup>4</sup> Universitat Pompeu Fabra, Barcelona, Spain, <sup>5</sup> Nanomalaria Group, Institute for Bioengineering of Catalonia, The Barcelona Institute of Science and Technology, Barcelona, Spain, <sup>6</sup> Barcelona Institute for Global Health, Barcelona Centre for International Health Research (Hospital Clínic – Universitat de Barcelona), Barcelona, Spain, <sup>7</sup> Institute of Nanoscience and Nanotechnology, University of Barcelona, Barcelona, Spain

Prions are a singular subset of proteins able to switch between a soluble conformation and a self-perpetuating amyloid state. Traditionally associated with neurodegenerative diseases, increasing evidence indicates that organisms exploit prion-like mechanisms for beneficial purposes. The ability to transit between conformations is encoded in the so-called prion domains, long disordered regions usually enriched in glutamine/asparagine residues. Interestingly, *Plasmodium falciparum*, the parasite that causes the most virulent form of malaria, is exceptionally rich in proteins bearing long Q/N-rich sequence stretches, accounting for roughly 30% of the proteome. This biased composition suggests that these protein regions might correspond to prion-like domains (PrLDs) and potentially form amyloid assemblies. To investigate this possibility, we performed a stringent computational survey for Q/N-rich PrLDs on *P. falciparum*. Our data indicate that ~10% of *P. falciparum* protein sequences have prionic signatures, and that this subproteome is enriched in regulatory proteins, such as transcription factors and RNA-binding proteins. Furthermore, we experimentally demonstrate for several of the identified PrLDs that, despite their disordered nature, they contain inner short sequences able to spontaneously self-assemble into amyloid-like structures. Although the ability of these sequences to nucleate the conformational conversion of the respective full-length proteins should still be demonstrated, our analysis suggests that, as previously described for other organisms, prion-like proteins might also play a functional role in *P. falciparum*.

**Keywords:** *Plasmodium*, protein aggregation, amyloid, prion, Q/N-rich sequences, protein disorder

## INTRODUCTION

Malaria caused approximately 445,000 deaths in 2016 and in the latest World Malaria Report (November 2017) the number of cases was estimated to be as many as 216 million. Although the global response to malaria is considered one of the world's great public health achievements, the spread of resistance against anti-malarial drugs and insecticides, has stalled the incidence and mortality decline since 2014.

*Plasmodium falciparum* is the species responsible for 85% of the malaria cases, causing the most severe form of the disease. The complete sequencing of *P. falciparum* genome has revealed some specific features that may shed light onto the biology and biochemistry of this deadly parasite (Gardner et al., 2002). A striking biased composition of its DNA was observed, with an overall AT content of 80.6%, a comparable AT enrichment only being observed in the social amoeba *Dictyostelium discoideum* (Eichinger et al., 2005). In *P. falciparum*, AT-rich codons present a significant preference toward encoding asparagines (N) over lysines, explaining why ~30% of its proteome is rich in long low complexity regions displaying an exceptional abundance of asparagine residues (Aravind et al., 2003; Singh et al., 2004).

Glutamine (Q)- and asparagine (N)-rich sequences have been shown to increase the propensity of a protein to form amyloids, and indeed the expansion of CAG trinucleotide repeats, encoding for Q, in different human proteins, results in developmental and neuromuscular disorders such as Huntington's disease, Kennedy disease, and several ataxias caused by the accumulation of intracellular protein aggregates in specific neuron types (Orr and Zoghbi, 2007; Williams and Paulson, 2008). Proteins with long N repeats have been shown to have an aggregation propensity even higher than poly-Q stretches (Tartaglia et al., 2005). Intriguingly, in spite of their inherent risk to promote aggregation, sequences with such amino acid compositions are common in yeast prions, and, thus are often referred to as prion domains (PrDs).

Among amyloids, prions are proteins with the unusual ability to adopt different structures and functional states, at least one of which is transmissible between individuals. In yeast, PrDs have been proved to be both sufficient and necessary for prion conformational conversion (Masison and Wickner, 1995). The detailed characterization of the prion phenomenon in yeast has provided important insights on the structural and sequential determinants of PrDs (Alberti et al., 2009). This knowledge has fuelled the development of computational algorithms able to identify prion like domains (PrLDs) in a genome-wide level in other organisms (Michelitsch and Weissman, 2000; Harrison and Gerstein, 2003; Espinosa Angarica et al., 2013, 2014; Lancaster et al., 2014; Zambrano et al., 2015), highlighting the existence of proteins bearing such intriguing sequences in all kingdoms of life (Espinosa Angarica et al., 2013, 2014).

It is now clear that evolution purges out proteins containing aggregation-prone regions, unless these sequences are needed for functional purposes (Monsellier and Chiti, 2007; Chen and Dokholyan, 2008). Given the intrinsic amyloid potential of PrLDs, their biological persistence suggests an evolutionary determination to maintain these regions. The word prion is usually associated with neurodegenerative diseases. However,

the recent identification of protein prion-like states executing physiological functions is rapidly changing this notation (Si, 2015). In higher eukaryotes, the earliest examples of functional prion-like polypeptides were described in *Aplysia* and *Drosophila*, where members of the CPEB protein family undergo prion conversion that facilitates memory formation (Heinrich and Lindquist, 2011; Majumdar et al., 2012). Cai et al. (2014) have revealed that the human proteins MAVS and ASC propagate respective downstream signals through prion conversion, and that this signal amplification strategy is crucial for the initiation of the innate immune response. More recently, non-pathogenic prion-like proteins have been described in plants and bacteria: luminidependens, an *Arabidopsis* protein, involved in flowering and plant memory regulation (Chakrabortee et al., 2016) and the transcription terminator Rho factor of the *Clostridium botulinum* pathogen (Pallarès et al., 2016; Yuan and Hochschild, 2017), respectively. These findings suggest that the conformational conversion and subsequent self-assembly that characterize prion-like proteins might be indeed an evolutionary conserved phenomenon (Maji et al., 2009; Tariq et al., 2013; Pallares and Ventura, 2017).

The enrichment of *P. falciparum* in N-rich low complexity sequences, soon suggested that this organism might contain a significant number of prion-like proteins, whose identification might contribute to understand its particular biology (Singh et al., 2004). Bioinformatic analysis found a correlation between the over-representation of homorepeat containing proteins and the abundance of proteins with putative PrLDs, which were proposed to account for as much as 25% of the parasite proteome (Singh et al., 2004). The biological significance of these protein domains is not clear (Muralidharan and Goldberg, 2013), but *P. falciparum* has evolved a very efficient proteostatic system to cope with such an aggregation-prone proteome (Muralidharan et al., 2012; Muralidharan and Goldberg, 2013; Przyborski et al., 2015).

In order to address the potential biological role for prion-like proteins in *P. falciparum*, we analyzed its proteome using a highly stringent computational approach, searching for the presence of Q/N-rich long regions displaying compositional similitude to *bona fide* prions (Toombs et al., 2010, 2012) and bearing specific amyloidogenic regions able to promote their self-assembly (Sabate et al., 2015b; Pallarès et al., 2016; Fernandez et al., 2017). This is the same approach that recently allowed us to propose the transcription terminator Rho factor as a first prion-like protein in bacteria (Pallarès et al., 2016; Pallares and Ventura, 2017). Applying this strategy, we have identified 503 PrLDs-containing proteins in *P. falciparum*, accounting for ~10% of its proteome. An analysis of the gene ontology (GO) terms enriched in this subproteome indicates that the proteins it contains might participate in important biological processes, such as regulation of gene expression or vesicle-mediated transport. Several of the functions assigned to the putative *P. falciparum* prion-like proteins are common to those reported in other eukaryotes, including humans, while other appear to be specific for this protozoan parasite.

Among all the identified prion-like candidates, we have selected three unrelated proteins and experimentally validated that their PrLDs contain specific short N-rich sequences able

to form amyloid fibrils; having thus the potential to trigger the conformational conversion of the *P. falciparum* proteins in which they are embedded.

Collectively, our study suggest that prion-like proteins may play a functional role in the complex parasite's biology.

## MATERIALS AND METHODS

### Dataset Construction

The *P. falciparum* (isolate 3D7) reference proteome (Proteome ID UP000001450, released 2015\_04, published on April 1, 2015) consisting of 5353 proteins was downloaded from UniProt (UniProt Consortium, 2015). This was first screened for the presence of Q/N-rich domains using an in-house developed Python script. Briefly, it scans for consecutive 80-residue windows and retrieves those with at least 30 Q/Ns. Once applied to *P. falciparum*'s proteome, it rendered 1300 proteins with at least one Q/N-rich domain. These were further scanned with PAPA (using default parameters) for intrinsically disordered regions and compositional bias resembling yeast prions, rendering 581 proteins. A final scan in search for soft amyloid cores within these PrLDs was performed using pWALTZ (using default parameters), resulting in a prion-like dataset of 503 proteins.

### Functional and Structural Enrichment Analysis

Gene ontology (Gene Ontology Consortium, 2015) terms (at the GO FAT category) and Pfam domain enrichment (Finn et al., 2016) were analyzed and clustered with the Functional Annotation Tool of DAVID 6.7 (Database for Annotation, Visualization and Integrated Discovery) (Huang da et al., 2009). The GO term clustering was performed with a high classification stringency and a  $p$ -value  $\leq 0.05$ . The Pfam list was obtained with a  $p$ -value  $\leq 0.05$  and final clustering was manually curated. From the 503 proteins in the prion-like proteins dataset, 487 were identified and processed by DAVID. The translation rates of the 10% highest scoring prion-like proteins at the different stages of *P. falciparum* life cycle were retrieved from Le Roch et al. (2003). For every protein entry, the developmental stage with the highest translation rate was considered.

### Peptide Prediction, Synthesis, and Preparation

The sequences of Sec24 (C0H489), the translation initiation factor-like protein IF2c (Q8IBA3) and the protein kinase PK4 (C6KTB8) were also analyzed with the PLAAC (Lancaster et al., 2014) algorithm. The resulting sequences, their position in the full-length protein and their scores are shown in **Figure 2**. The 21-residue peptides corresponding to the PrLDs predicted by pWALTZ were purchased from CASLO ApS (Scion Denmark Technical University). Peptide stock solutions were prepared solubilizing the lyophilized peptides at a final concentration of 5 mM in 100% dimethyl sulfoxide and stored at  $-80^{\circ}\text{C}$ . Before

each analysis, the samples were diluted to 150  $\mu\text{M}$  in phosphate-buffered saline (PBS) buffer. For aggregation assays, the diluted samples were incubated for 48 h at  $25^{\circ}\text{C}$ .

### Binding to Amyloid Dyes

The fluorescence spectra of the binding of 25  $\mu\text{M}$  Thioflavin-T (ThT) to peptide fibrils were recorded using a Cary Eclipse spectrofluorometer (Varian, Palo Alto, CA, United States) with an excitation wavelength of 440 nm and emission range from 460 to 600 nm at  $25^{\circ}\text{C}$  in PBS buffer. Peptides were equilibrated at room temperature for 2 min before the measurement and solutions without peptide were employed as negative controls. Excitation and emission slit widths of 10 nm were used. For the Thioflavin-S (ThS) staining assays, aggregated peptides were incubated for 1 h in the presence of 125  $\mu\text{M}$  ThS in PBS. Then, the samples were centrifuged ( $14,000 \times g$  for 5 min) and the precipitated fraction washed twice with PBS and finally placed on a microscope slide and sealed. Images of the aggregated peptides bound to ThS were obtained at 40-fold magnification under UV light or phase contrast in a Leica fluorescence microscope (Leica DMRB, Heidelberg, Germany).

### Secondary Structure Determination

Attenuated total reflectance Fourier transform infrared (ATR FT-IR) spectroscopy analysis of peptide fibrils were performed using a Bruker Tensor FT-IR Spectrometer (Bruker Optics, Berlin, Germany) with a Golden Gate MKII ATR accessory. Each spectrum consisted of 16 independent scans, measured at spectral resolution of  $1\text{ cm}^{-1}$  within the  $1800\text{--}1500\text{ cm}^{-1}$  range. All spectral data were acquired and normalized using the OPUS MIR Tensor 27 software. Infrared spectra between  $1725$  and  $1575\text{ cm}^{-1}$  were fitted through overlapping Gaussian curves, and the amplitude and area for each Gaussian function were calculated employing the non-linear peak-fitting program (PeakFit package, Systat Software, San Jose, CA, United States). Aggregated peptides were prepared at 150  $\mu\text{M}$  in PBS buffer and incubated for 48 h at  $25^{\circ}\text{C}$ . The PBS buffer without peptide was used as a control and subtracted from the absorbance signal before deconvolution.

### Transmission Electron Microscopy

Samples of aggregated peptides obtained as described previously, were placed onto carbon-coated copper grids and incubated for 5 min. The grids were washed with distilled water and negatively stained with 2% (w/v) uranyl acetate for 2 min. Micrographs were obtained in a JEM-1400 (JEOL, Japan) transmission electron microscope (TEM) operated at 80 kV accelerating voltage.

### In vivo Amyloid-Like Detection

Cultures of *P. falciparum* strain 3D7 were grown *in vitro* in group B human red blood cells (RBCs), purchased from the Banc de Sang i Teixits<sup>1</sup>, using previously described conditions (Cranmer et al., 1997). Briefly, parasites (thawed from glycerol stocks) were cultured at  $37^{\circ}\text{C}$  in T-Flasks containing RBCs in Roswell Park Memorial Institute (RPMI) complete medium under a gas

<sup>1</sup><http://www.bancsang.net>

mixture of 92% N<sub>2</sub>, 5% CO<sub>2</sub>, and 3% O<sub>2</sub>. Synchronized cultures were obtained by 5% sorbitol lysis (Lambros and Vanderberg, 1979) and the medium was changed every 2 days maintaining 3% hematocrit and a parasitemia below 5%. Staining with PROTEOSTAT<sup>®</sup> protein aggregation assay (Enzo Life Sciences, Inc.) was performed according to the manufacturer's instructions. Briefly, 200  $\mu$ l of *P. falciparum* culture were harvested and washed twice with 1 ml of 7.5 mg BSA/ml PBS (PBS/BSA); the resulting cell pellet was taken up in 200  $\mu$ l of PBS/BSA containing 2  $\mu$ g/ml Hoechst 33342 and PROTEOSTAT<sup>®</sup> (1:3000 stock dilution), and incubated for 30 min at room temperature in the dark before being washed again twice with 1 ml of PBS/BSA. 10  $\mu$ l of the washed cell suspension were transferred into a Lab-Tek chambered coverglass (Nunc, Thermo Fisher Scientific) containing 180  $\mu$ l of PBS/BSA and finally analyzed with a Leica TCS SP5 laser scanning confocal microscope, using a 63 $\times$  immersion oil objective with 1.4 numeric aperture. Hoechst 33342 and PROTEOSTAT<sup>®</sup> were detected, respectively, by excitation through 405 and 488 nm lasers. Emission was collected between 415 and 500 nm for Hoechst 33342, and between 590 and 670 for PROTEOSTAT<sup>®</sup>.

## RESULTS

### The *P. falciparum* Proteome Is Enriched in Proteins With PrLDs

The *P. falciparum* proteome contains an unusually high amount of low complexity regions; long domains enriched in certain amino acids and without a defined secondary structure. Low complexity regions are present in 30–90% of *P. falciparum* proteins, depending on the detection stringency (Singh et al., 2004; DePristo et al., 2006), and they are specially enriched in N residues. It has been proposed that these disordered protein regions might share certain properties with the classical yeast Q/N-rich PrLDs (Faux et al., 2005; Tompa and Fuxreiter, 2008; Fuxreiter, 2012; Fuxreiter and Tompa, 2012; Malinowska et al., 2013; Pallarès et al., 2016), and potentially support the formation of prion-like macromolecular assemblies (Singh et al., 2004; Espinosa Angarica et al., 2013).

In order to evaluate the presence of Q/N-rich prion-like proteins in *P. falciparum*, we examined its proteome combining detection of local Q/N-enrichment together with PAPA and pWALTZ predictions (Toombs et al., 2012; Zambrano et al., 2015; Pallarès et al., 2016) (see section “Materials and Methods”). Thus, any predicted PrLD in our subproteome would fulfill the following requirements: being Q/N-rich, disordered (PAPA includes the disorder predictor FoldIndex; Prilusky et al., 2005), compositionally similar to yeast PrLDs and contain a short sequence stretch able to facilitate its conversion into an amyloid-like state; we have generically named these stretches “soft amyloid cores,” because their amyloid propensity is significantly lower than the classical amyloid regions of pathogenic proteins, but still enough to promote protein self-assembly (Batlle et al., 2017).

The analysis shows that 1300 proteins (24.3% of the proteome) bear at least one Q/N-rich domain, in excellent agreement with previous studies (Singh et al., 2004). 581 of these Q/N-rich

domains (44.7%) also display an amino acid composition similar to that of yeast PrLDs and are disordered, as predicted by PAPA, and among them, 503 domains (86.6%) contain a soft amyloid core as predicted by pWALTZ (Supplementary Table S1). Overall, we conclude that 9.4% of the *P. falciparum* proteome may have the physicochemical properties and the aggregation potential to behave as a prion. This value is lower than previously estimated applying other computational approaches based only in Q/N richness (Singh et al., 2004) or than the one estimated using only compositional similitude to yeast PrLDs with PAPA, which predicts 22.5% of the parasite proteins as prion-like. The Q/N rich only dataset contains many long poly-N stretches, without any inner hydrophobic residue, a requisite to act as a prion (Toombs et al., 2010), whereas the PAPA only dataset include polypeptides unlikely to behave as prions *in vivo*, like membrane integral proteins, only because they display sequence stretches enriched in certain hydrophobic residues. In any case, even with the stringent approach used here, roughly 10% of the *P. falciparum* proteome seems to correspond to proteins displaying PrLDs and may thus have a high intrinsic aggregation propensity. This generates several important questions: Why *P. falciparum* proteins are so rich in PrLDs? Which are these putative prions? What are their roles in *P. falciparum*?

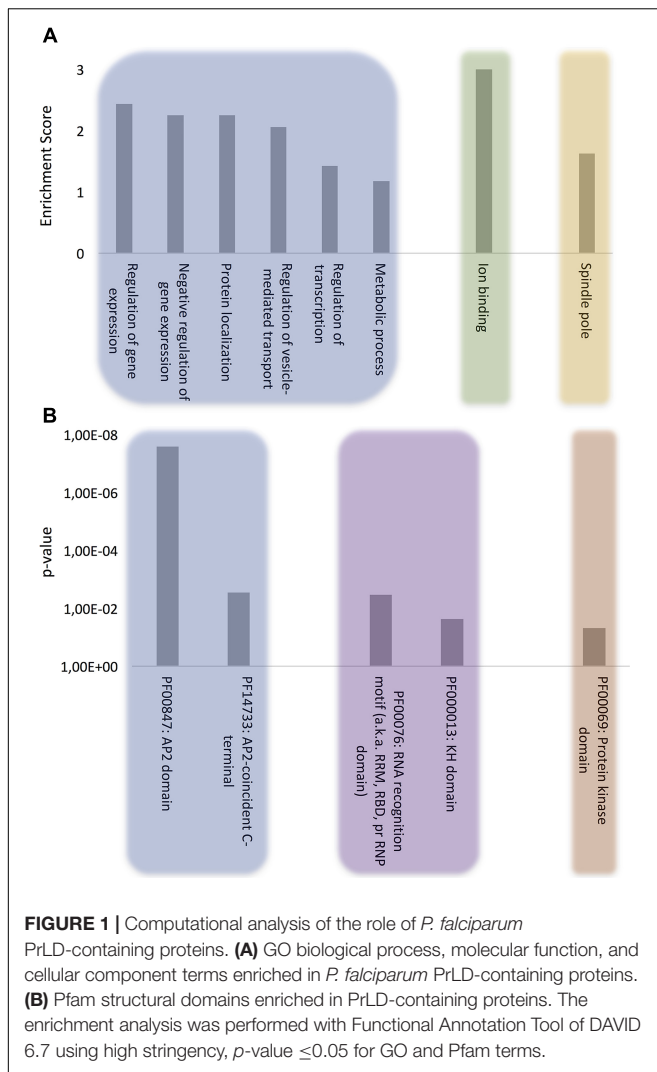
Previous works addressed these questions by analyzing the distribution of Q/N- or N-rich regions in the *P. falciparum* proteome (Singh et al., 2004). They detected such stretches in all protein families and all developmental stages of *P. falciparum*, without an evident association with any specific biological process. However, an increasing number of studies are uncovering PrLDs being connected to specific functions and processes in other species (Espinosa Angarica et al., 2014; Iglesias et al., 2015). We hypothesized that focusing the analysis in our curated subproteome may help to unravel the functional purpose, if any, of PrLDs in *P. falciparum*.

### Computational Analysis of the Role of Prion-Like Proteins in *P. falciparum*

The DAVID Functional Annotation Clustering Tool was employed to identify enriched GO categories (Huang et al., 2009; Gene Ontology Consortium, 2015) in the previously identified *P. falciparum* PrLD-containing proteins ( $p$ -value  $\leq 0.05$ ). It is worth to mention that this analysis is constrained by the fact that 60% of *P. falciparum* genes have unknown functions (Gardner et al., 2002), most of them have no clear homolog in other eukaryotes, and that the mechanisms underlying the main processes related to malaria pathogenesis in *P. falciparum* are still poorly understood. However, 487 out of the 503 proteins in our dataset were identified and processed by DAVID, which results in a coverage of 94.8% of our subproteome.

The proteins were clustered according to the following ontologies: biological process, molecular function, and cellular component (Figure 1A and Supplementary Table S2). The most significant biological process gene clusters include “regulation of gene expression,” “negative regulation of gene expression,” and “regulation of transcription.” The abundance of DNA and RNA binding proteins with PrLDs (Figure 1B) is consistent





with the observation that many of the prion-like proteins discovered initially in yeast (Alberti et al., 2009) and more recently in humans (King et al., 2012), plants (Chakrabortee et al., 2016) and bacteria proteomes (Iglesias et al., 2015) are proteins associated with gene expression and translation regulation such as transcription factors and RNA-binding proteins.

Other biological process terms could also be arranged into enriched clusters: “protein localization,” “regulation of vesicle-mediated transport,” and “metabolic process.” Importantly, the vesicle-mediated transport system and the trafficking of parasite proteins to diverse locations in the host cell are essential to promote new parasite phenotypes, playing a crucial role in host-pathogen interactions, as well as in disease pathogenesis and susceptibility (Miller et al., 2002; Hiller et al., 2004; Marti et al., 2005). Indeed, extracellular vesicles have been shown to act as delivery agents for prion-like proteins (Liu et al., 2017).

The analysis of molecular function domains in the set of *P. falciparum* prion-like proteins revealed that the only significantly enriched cluster was ion binding. A deeper analysis of the GO annotations indicates that ~33% of these proteins

function in DNA/RNA interaction and ~40% of them also contain structural domains related to nucleotide binding, such as Zinc fingers. In fact, the functions associated to nucleotide binding, especially RNA binding, appear to be associated to proteins containing PrLDs, regardless of the organism (Espinosa Angarica et al., 2014; Iglesias et al., 2015; Pallarès et al., 2016).

At this point, to dig a bit more on the functional role of our protein subset, we reanalyzed the MF category setting a  $p$ -value cutoff of 0.1. Several new GO terms came to light that could be grouped into two interesting MF subclusters: (i) chromatin remodeling, which is consistent with recent studies demonstrating that the physical properties of PrLDs can retarget critical chromatin regulatory complexes (Boulay et al., 2017) and facilitate heterochromatin assembly (Kataoka and Mochizuki, 2017) and (ii) GTPase regulatory activity, which is also detected in the PrLD-containing proteins of several other organisms (bacteria, plants, fungi, and invertebrates) (Espinosa Angarica et al., 2013); indeed, the canonical and best characterized yeast prion, Sup35, is a GTPase (Glover et al., 1997).

Analysis of the cellular component ontology category shows a specific enrichment at the spindle pole. In yeast, prion proteins have been shown to interact specifically with spindle pole proteins (Treusch and Lindquist, 2012) and spindle-associated proteins have been shown to be involved in self-assembly mediated phase separation in *Xenopus* (Jiang et al., 2015).

All the proteins in our Q/N-rich dataset have in common the presence of a disordered region of 80 amino acids in which at least the 37.5% of the residues (30/80) correspond to Q or N. This compositional similitude might imply a certain overlap of functions between proteins bearing PrLDs and those devoid of them. A GO analysis of Q/N-rich proteins without PrLDs (**Supplementary Figure S1**), shows that the molecular function and cellular location terms are different, but related, to those found in the Q/N-rich protein subset bearing PrLDs; nucleotide binding and cytoskeleton being the most enriched terms for these two categories, respectively. In contrast, Q/N-rich proteins without PrLDs are poorly represented in specific biological processes, being the most enriched one DNA repair (**Supplementary Figure S1**). This suggests that the compositional/sequential features of PrLDs might be important to specify the biological context in which the proteins act, whereas their generic molecular function depends mostly on the local enrichment in Q/N residues.

Prion-like proteins display a modular architecture in which one or several long and disordered PrLDs are adjacent to conventional globular domains and, accordingly, they tend to be large. We compared the average size of our protein subset with the one of the complete plasmodium proteome, confirming that proteins bearing PrLDs are effectively significantly longer (**Supplementary Figure S2**). To discard that the GO terms identified for PrLD-containing proteins would respond only to their differential size, we selected the subset of the largest 503 proteins in the proteome and performed a GO analysis. The resulting enriched terms did not coincide with those in our subset in any of the categories. The most enriched biological processes in large proteins were pathogenesis and single organismal cell-cell adhesion; the most enriched compartments were infected

host cell surface knob and host cell plasma membrane and the most enriched molecular functions were receptor activity and cell adhesion molecular binding.

In order to address if the expression of PrLD-containing proteins occurs preferentially at a given parasite stage, we analyzed the expression levels of the 10% top ranking proteins in our dataset at each of the different life cycle stages, as reported by Winzeler and colleagues (Le Roch et al., 2003). It turns out that, on the average, the highest translation rates for these proteins correspond to those at the merozoite and early ring stages (Supplementary Figure S3).

## Protein Domains in *P. falciparum* Prion-Like Proteins

To further evaluate the role of our collection of PrLD-containing proteins, we examined in detail their constituent functional domains (Finn et al., 2016) (Figure 1B). As expected, after clustering, the Pfam domains that were most often found in combination with PrLDs were involved in DNA/RNA binding, among which, the ApiAP2 stands out. The ApiAP2 family is homologous to the plant *Apetala2*/ethylene response factor (AP2/ERF) DNA-binding proteins, which comprise the second largest class of transcription factors in *A. thaliana*. Balaji et al. (2005) described that ApiAP2 proteins are likely to function as a family of apicomplexan parasite-specific transcription factors and that their amino acid sequences are highly conserved among orthologs. Strikingly, our data reveals that at least the 50% of the members composing this family in *P. falciparum* contain a PrLD. Several studies support their major role in mediating the regulation of stage-specific gene expression profiles in the parasite's development (Yuda et al., 2010; Painter et al., 2011; Modrzyńska et al., 2017) and suggest their crucial contribution to *P. falciparum* complexity and growth since very few ApiAP2 genes have been successfully knocked out (Behnke et al., 2010; Yuda et al., 2010).

The RNA recognition motif (RRM) is the most enriched RNA-binding domain (RBD) in our dataset. RRM is by far the most versatile and abundant RBDs, their fold being conserved from bacteria to higher eukaryotes (Reddy et al., 2015). This result is consistent with the observation that many of the human proteins with PrLDs contain an RRM motif and are involved in liquid-liquid phase transitions facilitating the formation of dynamic membraneless intracellular compartments, such as ribonucleoprotein (RNP) granules. They allow material exchange and fast assemblage and adaptation to different environments and cell states (Malinovska et al., 2013), the PrLDs in RNA binding proteins provide the special physicochemical properties that allow contacts between RNAs and proteins that sustain the liquid-like assemblies (Han et al., 2012; Kato et al., 2012). Indeed, the second most abundant RBD linked to our protein subproteome is the KH domain, a protein domain that was first identified in the human heterogeneous nuclear proteins (Siomi et al., 1993) and, together with RRM, constitutes the most abundant domain in RNA granules forming proteins (Kato et al., 2012). In *P. falciparum* RNPs are involved in

translation repression and posttranscriptional regulation of gene expression, critical for some stages of the parasite (Kramer, 2014).

The last enriched Pfam family includes the protein kinase domain. It is well-known that phosphorylation/dephosphorylation is the major control mechanism for many cellular functions. Consistently, recent studies carried out in *P. falciparum* reveal stage-specific profiles of protein phosphorylation, suggesting that reversible protein phosphorylation plays a key role in the regulation of the *Plasmodium* life cycle (Wu et al., 2009; Pease et al., 2013). So far, no PrLD-containing protein kinase has been characterized experimentally, but it is obvious that the ability to control the activity of these enzymes by modulating their assembly would have important physiological consequences.

## Predicted PrLD Soft Amyloid Cores in *P. falciparum* Proteins

Based on the above computational results, we selected three PrLD-containing proteins for their experimental characterization: the putative transport protein Sec24 (C0H489), the translation initiation factor-like protein IF2c (Q8IBA3), and the protein kinase PK4 (C6KTB8). These proteins are associated with functions (nucleotide binding, Q8IBA3), cellular components (vesicle mediated-transport, C0H489) and structural domains (kinase, C6KTB8) that are enriched in our dataset. The selected candidates have no functional or sequential relationship and have not been previously suggested to act as prions.

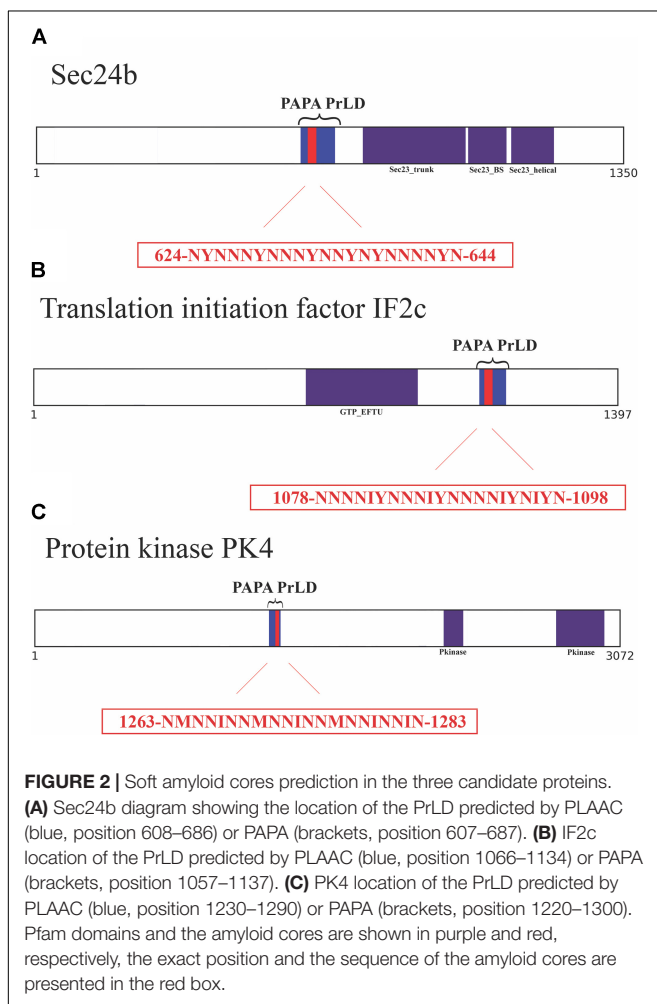
As a first candidate, we chose Sec24b, a member of the Sec24 family. Within this family, Sec24b is by far the most enriched in Q/N, which constitute 21.5% of the amino acids in the complete sequence and 51.9% of the PrLD. Sec24b is closely related to the mammalian Sec24C/D family and the yeast Sec24 homolog Lst1 (Lee et al., 2008). These proteins play a key role in shaping the vesicle, as well as in cargo selection and concentration (Roberg et al., 1999). They have a scaffolding function required to generate vesicles that can accommodate difficult cargo proteins, including large oligomeric assemblies.

As a second candidate, we selected IF2, one of the essential components for the initiation of protein synthesis. IF2 is a translation initiator factor acting as a GTPase that recruits the charged fMet-initiator tRNA onto the 30S ribosomal initiation complex (Antoun et al., 2003). From the three IF2 homologs described in *P. falciparum* (Haider et al., 2015), IF2c is the only one that holds a PrLD. It is worth to note that the IF2c C-terminal domain, where the PrLD maps, has the largest identity with bacterial IF2, a family of translation initiation factors rich in putative PrLDs (Iglesias et al., 2015). 20.8% of IF2c residues are Q/N and this proportion raises to 49.4% in the PrLD.

The third selected protein was the kinase PK4, a protein that is essential for completion of the blood stage of the disease (Zhang et al., 2012). Thus PK4 has been suggested as a novel target for the next generation of antimalarial compounds (Kahrstrom, 2012). PK4 phosphorylates eIF2 $\alpha$

and arrests global protein synthesis in schizonts (mature form of the blood cycle) and gametocytes (sexual form that infects the mosquitoes). 17.7% of PK4 amino acids are Q or N, with 55.5% of its PrLD corresponding to these polar residues.

To further confirm the presence of PrLDs in these proteins, and to define more precisely their boundaries, we used PLAAC (Lancaster et al., 2014), yet another composition-based predictor, in which, in contrast to PAPA, the length of the predicted PrLD also depends on the protein composition. PLAAC detected PrLDs overlapping with the regions previously identified by PAPA, in the three polypeptides (Figure 2).



We analyzed these three putative prion-like proteins using the same computational approach, we employed previously to detect and validate the soft amyloid cores present in *bona fide* yeast prions (Sant'Anna et al., 2016), in the pathogenic bacteria *C. botulinum* (Pallarès et al., 2016) and in human prion-like proteins (Batlle et al., 2017). The predicted soft amyloid cores for these *P. falciparum* proteins are shown in Table 1. Not surprisingly, these stretches are highly enriched in N residues, all containing >60% of this polar residue.

Interestingly enough, well-validated aggregation predictors like Aggrescan, Tango, and Zyggregator (Fernandez-Escamilla et al., 2004; Conchillo-Sole et al., 2007; Tartaglia and Vendruscolo, 2008), all failed to classify these stretches as aggregation-prone (Supplementary Table S3), likely because of their much lower hydrophobicity, when compared with the classical amyloid stretches present in pathogenic amyloidogenic proteins (Supplementary Table S4). One of the restraints in our prediction scheme is that the identified PrLD should be essentially disordered, as predicted with FoldIndex (Prilusky et al., 2005). In this structural context, the identified soft amyloid cores will be mostly exposed to solvent and able to establish intermolecular contacts, if they have this ability. Orthogonal analysis with alternative disorder prediction algorithms confirms that this is likely the case for the three proteins herein (Supplementary Table S5).

We synthesized 21-residue-long peptides corresponding to the detected soft amyloid cores and characterized their amyloid properties experimentally.

## Predicted PrLDs Soft Amyloid Cores Assemble Into $\beta$ -Sheet Rich Structures

As a first evaluation of the assembling properties of the selected peptides, we measured their ability to adopt a  $\beta$ -sheet-enriched structure, a hallmark of amyloid fibril formation (Nelson et al., 2005). To this aim, the peptides were prepared at 150  $\mu$ M in PBS and incubated during 48 h at 25°C. We used FT-IR spectroscopy and recorded the amide I region of the spectrum (1700–1600  $\text{cm}^{-1}$ ) (Figure 3 and Supplementary Table S6). This spectral region corresponds to the absorption of the carbonyl peptide bond group of the protein main chain and is sensitive to the protein conformation. Deconvolution of the spectra allowed us to assign the secondary structure elements and their relative contribution to the main absorbance. In the three cases, the main peaks mapped in the 1620–1630  $\text{cm}^{-1}$  region of the spectra, accounting for 50% or more of the absorbance signals, indicating that the peptides have acquired significant intermolecular  $\beta$ -sheet structure. Interestingly, no anti-parallel  $\beta$ -sheet band was detected ( $\sim 1690 \text{ cm}^{-1}$ ) in any of the samples; thus, suggesting that the detected  $\beta$ -strands in the self-assembled peptides would adopt preferentially a parallel disposition. The other detected signals were associated with the presence of disordered structure and turns (Figure 3 and Supplementary Table S6).

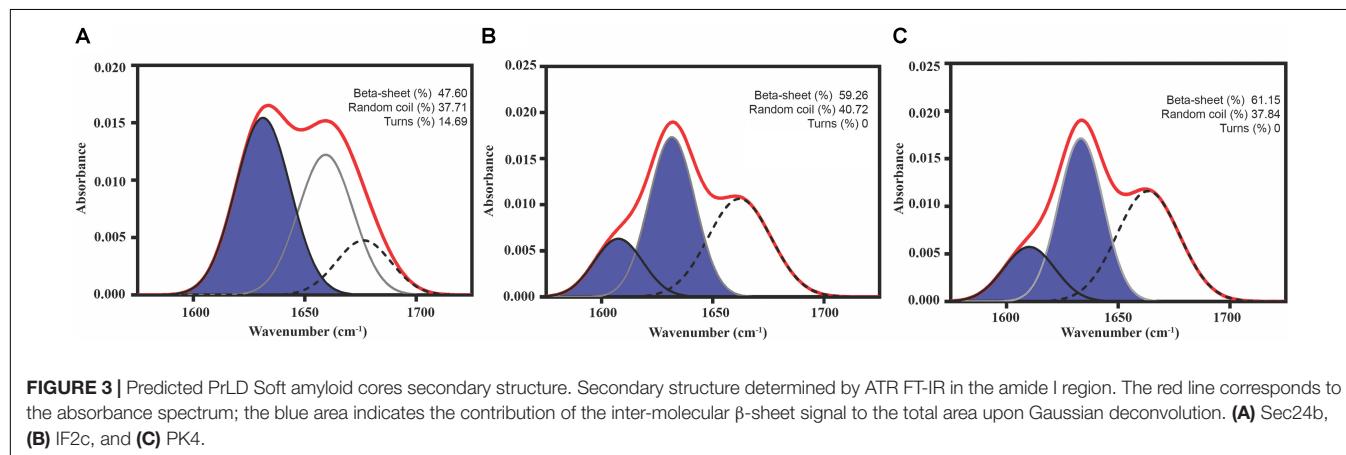
## Predicted PrLD Soft Amyloid Cores Form Amyloid-Like Fibrillar Structures

To assess if the identified  $\beta$ -sheet-rich assemblies correspond to amyloid-like structures, we used the amyloid-specific dyes ThT and ThS. After incubation at a concentration of 150  $\mu$ M in PBS during 48 h at 25°C, all the peptides were able to promote a large increase in the intensity of ThT fluorescence emission (Figure 4). In addition, areas rich in fibrous material were stained by ThS to yield a bright green-yellow fluorescence against a dark background (Figure 4).

**TABLE 1** | Predicted PrLD soft amyloid cores.

Protein	UniProt ID	PrLD amyloid core	N (%)	pWALTZ score	PAPA score	PLAAC score
Sec24b	C0H489	624-NYNNNNNNNNNNNNNNNNNN-644	71	84.62	0.20	45.28
IF2c	Q8IBA3	1078-NNNNIYNNNNNNNNIYNN-1098	62	87.71	0.07	36.18
PK4	C6KT8	1263-NMNNNNMNNNNMNNNN-1283	67	77.34	0.25	47.97

For each PrLD-containing protein, it is shown the UniProt ID, the 21 residue-long soft amyloid core with its respective position in the sequence, pWALTZ score, PAPA score, and PLAAC score with a cutoff of 73.55, 0.05, and >0, respectively. PAPA and PLAAC search for compositional similarity to yeast prion domains, defining the predicted PrLD, while pWALTZ scans for soft amyloid cores within them.



Transmission electron microscopy examination of the morphological features of the incubated peptide solutions (Figure 5) revealed that they effectively assemble into supramolecular structures. Sec24b formed long and straight fibrils, whereas IF2c and PK4 formed short and curly fibrillar structures.

Overall, biophysical analysis of the three predicted peptides demonstrates the ability of the candidate *P. falciparum* soft amyloid cores to nucleate the formation of  $\beta$ -sheet-rich amyloid-like structures.

## Detection of Intracellular Protein Aggregates in *P. falciparum*

The above experimental data suggest that ~10% of the *P. falciparum* proteome might possess the ability to establish amyloid-like contacts, at least transiently, *in vivo*, and thus; that at any time, a significant number of proteins might potentially aggregate in the parasite. We employed a permeable amyloid-specific dye (PROTEOSTAT®) to track the *in vivo* presence of intracellular amyloid-like aggregates in *P. falciparum*.

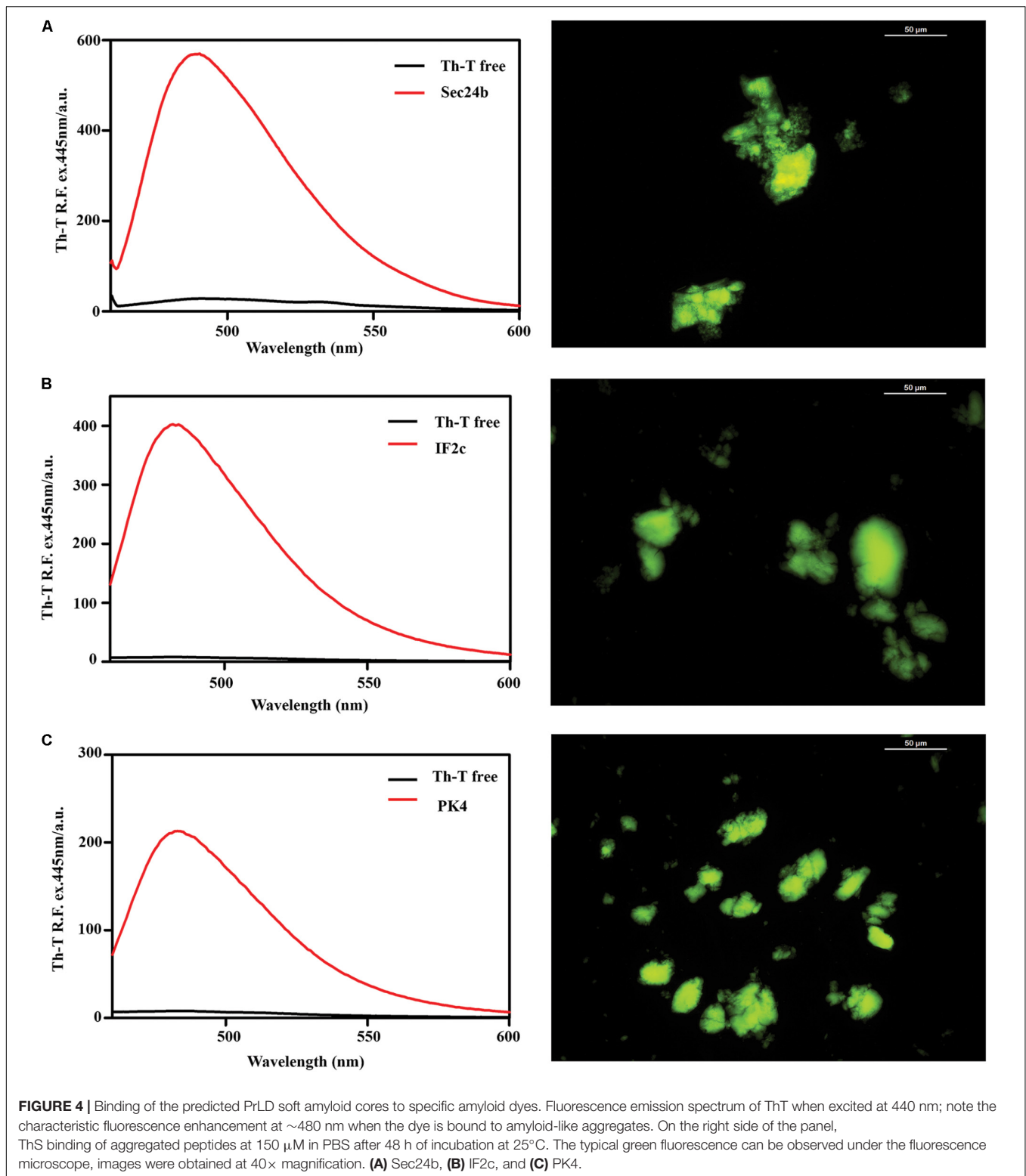
*Plasmodium falciparum* was grown in RBCs and then the culture was incubated with the amyloid dye. We observed colocalization between PROTEOSTAT® fluorescence and the cytosol of *P. falciparum*-infected RBCs, whose nuclei stained with Hoechst 33342. The images evidenced the lack of structures able to bind the dye in non-infected erythrocytes, and that, accordingly, only upon infection by *Plasmodium*, red fluorescent amyloid foci are evident inside parasitized RBCs (Figure 6), demonstrating the high amyloid load that this parasite supports at this stage.

## DISCUSSION

Many lines of evidence suggest that prion-like proteins can be both harmful and beneficial for the cell. The propensity of a protein to behave as a prion is encoded in its amino acid sequence (Sabate et al., 2015a). In particular, long and disordered N/Q-rich sequences seem to facilitate conformational conversion into functional amyloid-like states (Alberti et al., 2009). The occurrence of N/Q-rich sequence stretches varies substantially between organisms, with *P. falciparum* having one of the most enriched proteomes in this kind of regions, and specifically in N-rich sequences (Aravind et al., 2003). Accordingly, it has been assumed that prion-like proteins would be common in this organism (Michelitsch and Weissman, 2000). Long Q- and N-homorepeats are inherently aggregation-prone (Halfmann et al., 2011). However, these sequence stretches alone are not sufficient to sustain a prion-like behavior (Toombs et al., 2010). Here, using a stringent computational approach that considers that PrLDs should not be only Q/N-rich, but also display compositional similitude to *bona fide* yeast PrDs and encode for at least one specific short sequence stretch with moderate, but significant, amyloid propensity (Sabate et al., 2015b), we concluded that 503 polypeptides in *P. falciparum* fulfill the requirements to potentially behave as prion-like proteins. This accounts for ~10% of the proteome, which despite being a lower fraction than previously proposed (~25%) (Singh et al., 2004), still might constitute a high prionic load for the parasite.

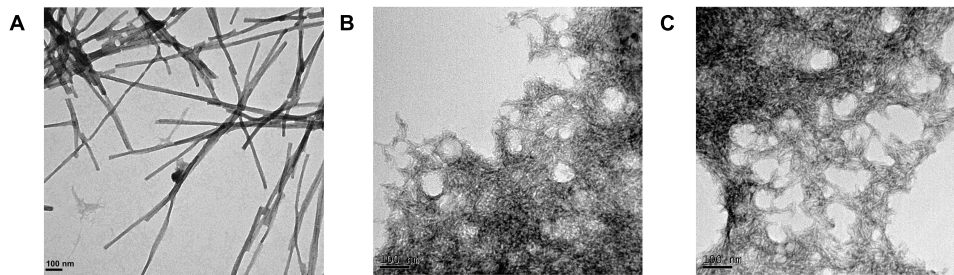
*A priori*, the presence of PrLDs might be dangerous for *Plasmodium*, since prion-like proteins have an intrinsic propensity to aggregate and, in humans, disease-linked mutations occur preferentially in the PrLDs of these polypeptides



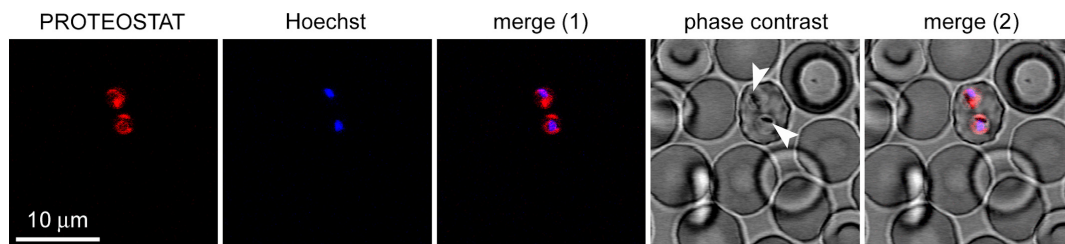


(Kim et al., 2013). On the other hand, these PrLDs might act as conformational switches that control protein assembly and thus protein function to allow adaptation to the changing environment that *P. falciparum* faces during its life cycle.

Importantly, the PrLD-containing subproteome, we identify here is associated with defined domains and functionalities in the parasite, which suggests that Q/N-rich PrLDs do not occur randomly in the *P. falciparum* proteome. This assumption is



**FIGURE 5** | Fibrillar structures formed by the predicted PrLD soft amyloid cores. Representative TEM images for **(A)** Sec24b, **(B)** IF2c, and **(C)** PK4 aggregated peptides at 150  $\mu$ M in PBS after 48 h of incubation at 25°C.



**FIGURE 6** | Fluorescence microscopy analysis of the presence of protein aggregates in *P. falciparum*-infected RBCs (pRBCs). The selected field shows a single pRBC in early trophozoite stage, indicated by its characteristic nuclear Hoechst blue fluorescence among enucleated non-parasitized erythrocytes. The amyloid-specific dye PROTEOSTAT<sup>®</sup> reveals protein aggregates in the cytosol of the two parasite cells present in the pRBC. The arrowheads indicate nascent hemozoin crystals in the food vacuole of *Plasmodium*.

supported by the fact that PrLDs are associated with similar GO-clusters in organisms as divergent as *Plasmodium*, yeast, *Dictyostelium*, *Drosophila* and humans (Malinovska et al., 2013, 2015). For instance, the role of PrLDs-containing proteins in DNA and RNA binding is well conserved, with the RRM domain being among the most enriched PrLD-associated domains in these organisms. Indeed, 25% of the *P. falciparum* proteins bearing an RRM domain also contain a predicted PrLD. We found that this domain association is also conserved in *Plasmodium vinckei* and *Plasmodium yoelii*, with 13 and 15% of RRM-containing proteins having a Q/N-rich PrLD.

*Plasmodium falciparum* PrLDs exhibit specific associations with domains and functions not detected in other organisms, such as the ApiAP2 proteins, with 50% of their members displaying a PrLD. These proteins have been postulated as the main transcriptional regulators in *Plasmodium* parasites and the other Apicomplexa (Balaji et al., 2005). Importantly, according to our analysis, the presence of PrLDs within AP2 transcription factors also seems to be evolutionary, with 39, 29, and 24% of the AP2 proteins in *P. vinckei*, *P. yoelii*, and *Plasmodium berghei* displaying Q/N-rich PrLDs, respectively. The association between PrLDs and the regulation of vesicle-mediated transport is also a specific feature of *Plasmodium*. This process allows the trafficking of some parasite proteins to the erythrocyte membrane (Miller et al., 2002; Hiller et al., 2004; Marti et al., 2005).

*Plasmodium* is an obligate parasite that has evolved to survive in different hosts and cell types. It has a complex life cycle with cellular stages that differ in shape, size, metabolic

activity, and resource requirements. Hence, to sustain this complexity, *Plasmodium* requires an efficient regulation, to which the conformational conversion of regulatory proteins bearing PrLDs might contribute. Changes in local protein concentration, binding to nucleic acids and posttranslational modifications have been shown to modulate the assembly of PrLDs, the functional outcome depending on the particular assembled protein (Alberti, 2017).

We and others have suggested that certain short amyloidogenic sequence stretches embedded in PrLDs contribute significantly to prion formation, maintenance, and transmission, at least in yeast (Crow et al., 2011; Sabate et al., 2015a; Sant'Anna et al., 2016; Batlle et al., 2017). The computational search for such regions in the putative PrLDs of a large number of bacterial proteomes (Iglesias et al., 2015), previously thought to lack prions, and a subsequent experimental validation, allowed us to propose that the Rho transcription terminator might constitute a first bacterial prion (Pallarès et al., 2016; Pallares and Ventura, 2017). Soon after, Yuan and Hochschild (2017) confirmed the ability of this protein to adopt an infectious state, leading to global changes in the transcriptome. Here, we used the same approach to study the amyloidogenic potential of three *P. falciparum* PrLD-containing proteins: the translation initiation factor 2c, the kinase PK4, both involved in gene expression regulation (Zhang et al., 2012; Haider et al., 2015) and Sec24b, involved in vesicle trafficking (Lee et al., 2008). Our data provide compelling evidence that, *in vitro*, all three candidate proteins contain short nucleating regions

embedded in the PrLDs able to spontaneously self-assemble into amyloid-like structures. The presence of such stretches does not necessarily imply that the correspondent large full-length proteins would behave in a prion-like manner, and this behavior should be experimentally validated. However, several indirect evidences suggest that this could be the case: (i) we have shown that when a predicted short amyloid sequence is administered to cells in its amyloid state it is able to seed the conformational conversion of the complete endogenous protein and its subsequent aggregation into a prionic form (Sant'Anna et al., 2016), (ii) the amyloid stretch, we identified in the PrLD of the bacterial Rho terminator factor has been shown to be absolutely essential for its self-assembly and prion activity (Yuan and Hochschild, 2017), and (iii) Vorberg and colleagues have shown that the amyloid sequence, we pinpointed in the PrLD of a model prion protein is the only region required for the induction, propagation, and inheritance of the prion state in the mammalian cytosol (Duernberger et al., 2018).

Overall, we identified a subset of putative Q/N-rich prion-like proteins in *P. falciparum* associated with specific biological processes and experimentally validated that their highly polar and disordered PrLDs contain cryptic sequences able to self-assemble into amyloids. The structural characterization and *in vivo* validation of the properties of the identified proteins is challenging, but it is worth the effort, since it might uncover a first *bona fide* prion in *Plasmodium*.

## REFERENCES

- Alberti, S. (2017). Phase separation in biology. *Curr. Biol.* 27, R1097–R1102. doi: 10.1016/j.cub.2017.08.069
- Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158. doi: 10.1016/j.cell.2009.02.044
- Antoun, A., Pavlov, M. Y., Andersson, K., Tenson, T., and Ehrenberg, M. (2003). The roles of initiation factor 2 and guanosine triphosphate in initiation of protein synthesis. *EMBO J.* 22, 5593–5601. doi: 10.1093/emboj/cdg525
- Aravind, L., Iyer, L. M., Welles, T. E., and Miller, L. H. (2003). *Plasmodium* biology: genomic gleanings. *Cell* 115, 771–785. doi: 10.1016/S0092-8674(03)01023-7
- Balaji, S., Babu, M. M., Iyer, L. M., and Aravind, L. (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* 33, 3994–4006. doi: 10.1093/nar/gki709
- Battle, C., de Groot, N. S., Iglesias, V., Navarro, S., and Ventura, S. (2017). Characterization of soft amyloid cores in human prion-like proteins. *Sci. Rep.* 7:12134. doi: 10.1038/s41598-017-09714-z
- Behnke, M. S., Wootton, J. C., Lehmann, M. M., Radke, J. B., Lucas, O., Nawas, J., et al. (2010). Coordinated progression through two subtranscriptomes underlies the tachyzoite cycle of *Toxoplasma gondii*. *PLoS One* 5:e12354. doi: 10.1371/journal.pone.0012354
- Boulay, G., Sandoval, G. J., Riggi, N., Iyer, S., Buisson, R., Naigles, B., et al. (2017). Cancer-specific retargeting of BAF complexes by a prion-like domain. *Cell* 171, 163.e19–178.e19. doi: 10.1016/j.cell.2017.07.036
- Cai, X., Chen, J., Xu, H., Liu, S., Jiang, Q. X., Halfmann, R., et al. (2014). Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation. *Cell* 156, 1207–1222. doi: 10.1016/j.cell.2014.01.063

## AUTHOR CONTRIBUTIONS

IP, XF-B, and SV conceived the experiments. IP, NG, VI, RS, and AB conducted the experiments. IP, NG, VI, XF-B, and SV analyzed the results. IP, NG, VI, and AB prepared the figures. IP, NG, and SV wrote the main manuscript text. All authors reviewed the manuscript.

## FUNDING

This work was funded by grants BIO2016-78310-R to SV and BIO2014-52872-R to XF-B (Ministerio de Economía y Competitividad, Spain), which included FEDER funds, and by ICREA, ICREA ACADEMIA 2015 to SV. ISGlobal and IBEC are members of the CERCA Program, Generalitat de Catalunya.

## ACKNOWLEDGMENTS

Electron microscopy experiments were performed in the UAB Microscopy facility.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.01737/full#supplementary-material>

- Chakrabortee, S., Kayatekin, C., Nweby, G. A., Mendillo, M. L., Lancaster, A., and Lindquist, S. (2016). Luminidependens (LD) is an Arabidopsis protein with prion behavior. *Proc. Natl. Acad. Sci. U.S.A.* 113, 6065–6070. doi: 10.1073/pnas.1604478113
- Chen, Y., and Dokholyan, N. V. (2008). Natural selection against protein aggregation on self-interacting and essential proteins in yeast, fly, and worm. *Mol. Biol. Evol.* 25, 1530–1533. doi: 10.1093/molbev/msn122
- Conchillo-Sole, O., de Groot, N. S., Aviles, F. X., Vendrell, J., Daura, X., and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* 8:65. doi: 10.1186/1471-2105-8-65
- Cranmer, S. L., Magowan, C., Liang, J., Coppel, R. L., and Cooke, B. M. (1997). An alternative to serum for cultivation of *Plasmodium falciparum* in vitro. *Trans. R. Soc. Trop. Med. Hyg.* 91, 363–365. doi: 10.1016/S0035-9203(97)90110-3
- Crow, E. T., Du, Z., and Li, L. (2011). A small, glutamine-free domain propagates the [SWT<sup>+</sup>] prion in budding yeast. *Mol. Cell. Biol.* 31, 3436–3444. doi: 10.1128/MCB.05338-11
- DePristo, M. A., Zilversmit, M. M., and Hartl, D. L. (2006). On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene* 378, 19–30. doi: 10.1016/j.gene.2006.03.023
- Duernberger, Y., Liu, S., Riemschoss, K., Paulsen, L., Bester, R., Kuhn, P. H., et al. (2018). Prion replication in the mammalian cytosol: functional regions within a prion domain driving induction, propagation and inheritance. *Mol. Cell. Biol.* doi: 10.1128/MCB.00111-18 [Epub ahead of print].
- Eichinger, L., Pacheb, J. A., Glockner, G., Rajandream, M. M., Suggang, R., Berriman, M., et al. (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435, 43–57. doi: 10.1038/nature03481

- Espinosa Angarica, V., Angulo, A., Giner, A., Losilla, G., Ventura, S., and Sancho, J. (2014). PrionScan: an online database of predicted prion domains in complete proteomes. *BMC Genomics* 15:102. doi: 10.1186/1471-2164-15-102
- Espinosa Angarica, V., Ventura, S., and Sancho, J. (2013). Discovering putative prion sequences in complete proteomes using probabilistic representations of Q/N-rich domains. *BMC Genomics* 14:316. doi: 10.1186/1471-2164-14-316
- Faux, N. G., Bottomley, S. P., Lesk, A. M., Irving, J. A., Morrison, J. R., de la Banda, M. G., et al. (2005). Functional insights from the distribution and role of homopeptide repeat-containing proteins. *Genome Res.* 15, 537–551. doi: 10.1101/gr.3096505
- Fernandez, M. R., Batlle, C., Gil-Garcia, M., and Ventura, S. (2017). Amyloid cores in prion domains: key regulators for prion conformational conversion. *Prion* 11, 31–39. doi: 10.1080/19336896.2017.1282020
- Fernandez-Escamilla, A. M., Rouseau, F., Schymkowitz, J., and Serrano, L. (2004). Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22, 1302–1306. doi: 10.1038/nbt.1012
- Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., et al. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* 44, D279–D285. doi: 10.1093/nar/gkv1344
- Fuxreiter, M. (2012). Fuzziness: linking regulation to protein dynamics. *Mol. Biosyst.* 8, 168–177. doi: 10.1039/c1mb05234a
- Fuxreiter, M., and Tompa, P. (2012). Fuzzy complexes: a more stochastic view of protein function. *Adv. Exp. Med. Biol.* 725, 1–14. doi: 10.1007/978-1-4614-0659-4\_1
- Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511. doi: 10.1038/nature01097
- Gene Ontology Consortium (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res.* 43, D1049–D1056. doi: 10.1093/nar/gku1179
- Glover, J. R., Kowal, A. S., Schirmer, E. C., Patino, M. M., Liu, J. J., and Lindquist, S. (1997). Self-seeded fibers formed by Sup35, the protein determinant of [PSI<sup>+</sup>], a heritable prion-like factor of *S. cerevisiae*. *Cell* 89, 811–819. doi: 10.1016/S0092-8674(00)80264-0
- Haider, A., Allen, S. M., Jackson, K. E., Ralph, S. A., and Habib, S. (2015). Targeting and function of proteins mediating translation initiation in organelles of *Plasmodium falciparum*. *Mol. Microbiol.* 96, 796–814. doi: 10.1111/mmi.12972
- Halfmann, R., Alberti, S., Krishnan, R., Lyle, N., O'Donnell, C. W., King, O. D., et al. (2011). Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins. *Mol. Cell* 43, 72–84. doi: 10.1016/j.molcel.2011.05.013
- Han, T. W., Kato, M., Xie, S., Wu, L. C., Mirzaei, H., Pei, J., et al. (2012). Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell* 149, 768–779. doi: 10.1016/j.cell.2012.04.016
- Harrison, P. M., and Gerstein, M. (2003). A method to assess compositional bias in biological sequences and its application to prion-like glutamine/asparagine-rich domains in eukaryotic proteomes. *Genome Biol.* 4:R40. doi: 10.1186/gb-2003-4-6-r40
- Heinrich, S. U., and Lindquist, S. (2011). Protein-only mechanism induces self-perpetuating changes in the activity of neuronal *Aplysia* cytoplasmic polyadenylation element binding protein (CPEB). *Proc. Natl. Acad. Sci. U.S.A.* 108, 2999–3004. doi: 10.1073/pnas.1019368108
- Hiller, N. L., Bhattacharjee, S., van Ooij, C., Liolio, K., Harrison, T., Lopez-estrano, C., et al. (2004). A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 306, 1934–1937. doi: 10.1126/science.1102737
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211
- Iglesias, V., de Groot, N. S., and Ventura, S. (2015). Computational analysis of candidate prion-like proteins in bacteria and their role. *Front. Microbiol.* 6:1123. doi: 10.3389/fmicb.2015.01123
- Jiang, H., Wang, S., Huang, Y., He, X., Cui, H., Zhu, X., et al. (2015). Phase transition of spindle-associated protein regulate spindle apparatus assembly. *Cell* 163, 108–122. doi: 10.1016/j.cell.2015.08.010
- Kahrstrom, C. T. (2012). Parasite physiology: *Plasmodium* gets the PK4 blood test. *Nat. Rev. Microbiol.* 10:237. doi: 10.1038/nrmicro2774
- Kataoka, K., and Mochizuki, K. (2017). Heterochromatin aggregation during DNA elimination in *Tetrahymena* is facilitated by a prion-like protein. *J. Cell Sci.* 130, 480–489. doi: 10.1242/jcs.195503
- Kato, M., Han, T. M., Xie, S., Shi, K., Du, X., Wu, L. C., et al. (2012). Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753–767. doi: 10.1016/j.cell.2012.04.017
- Kim, H. J., Kim, N. C., Wang, Y. D., Scarborough, E. A., Moore, J., Diaz, Z., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467–473. doi: 10.1038/nature11922
- King, O. D., Gitler, A. D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* 1462, 61–80. doi: 10.1016/j.brainres.2012.01.016
- Kramer, S. (2014). RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. *Wiley Interdiscip. Rev. RNA* 5, 263–284. doi: 10.1002/wrna.1207
- Lambros, C., and Vanderberg, J. P. (1979). Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J. Parasitol.* 65, 418–420. doi: 10.2307/3280287
- Lancaster, A. K., Nutter-Upham, A., Lindquist, S., and King, O. D. (2014). PLAAC: a web and command-line application to identify proteins with Prion-Like Amino Acid Composition. *Bioinformatics* 30, 2–3. doi: 10.1093/bioinformatics/btu310
- Le Roch, K. G., Zhou, Y., Blair, P. L., Grainger, M., Moch, J. K., Haynes, J. D., et al. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301, 1503–1508. doi: 10.1126/science.1087025
- Lee, M. C., Moura, P. A., Miller, E. A., and Fidock, D. A. (2008). *Plasmodium falciparum* Sec24 marks transitional ER that exports a model cargo via a diacidic motif. *Mol. Microbiol.* 68, 1535–1546. doi: 10.1111/j.1365-2958.2008.06250.x
- Liu, S., Hossinger, A., Gobbels, S., and Vorberg, I. M. (2017). Prions on the run: how extracellular vesicles serve as delivery vehicles for self-templating protein aggregates. *Prion* 11, 98–112. doi: 10.1080/19336896.2017.1306162
- Maji, S. K., Perrin, M. H., Sawaya, M. R., Jessberger, S., Vadodaria, K., Rissman, R. A., et al. (2009). Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325, 328–332. doi: 10.1126/science.1173155
- Majumdar, A., Cesario, W. C., White-Grindley, E., Jiang, H., Ren, F., Khan, M. R., et al. (2012). Critical role of amyloid-like oligomers of *Drosophila* Orb2 in the persistence of memory. *Cell* 148, 515–529. doi: 10.1016/j.cell.2012.01.004
- Malinowska, L., Kroschwald, S., and Alberti, S. (2013). Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim. Biophys. Acta* 1834, 918–931. doi: 10.1016/j.bbapap.2013.01.003
- Malinowska, L., Palm, S., Gibson, K., Verbavatz, J. M., and Alberti, S. (2015). *Dictyostelium discoideum* has a highly Q/N-rich proteome and shows an unusual resilience to protein aggregation. *Proc. Natl. Acad. Sci. U.S.A.* 112, E2620–E2629. doi: 10.1073/pnas.1504459112
- Marti, M., Baum, J., Rug, M., Tilley, L., and Cowman, A. F. (2005). Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. *J. Cell Biol.* 171, 587–592. doi: 10.1083/jcb.200508051
- Masison, D. C., and Wickner, R. B. (1995). Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. *Science* 270, 93–95. doi: 10.1126/science.270.5233.93
- Michelitsch, M. D., and Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11910–11915. doi: 10.1073/pnas.97.22.11910
- Miller, L. H., Baruch, D. I., Marsh, K., and Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature* 415, 673–679. doi: 10.1038/415673a
- Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., et al. (2017). A knockout screen of ApiAP2 genes reveals networks of interacting transcriptional regulators controlling the *Plasmodium* life cycle. *Cell Host Microbe* 21, 11–22. doi: 10.1016/j.chom.2016.12.003
- Monsellier, E., and Chiti, F. (2007). Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep.* 8, 737–742. doi: 10.1038/sj.embor.7401034
- Muralidharan, V., and Goldberg, D. E. (2013). Asparagine repeats in *Plasmodium falciparum* proteins: good for nothing? *PLoS Pathog.* 9:e1003488. doi: 10.1371/journal.ppat.1003488



- Muralidharan, V., Oksman, A., Pal, P., Lindquist, S., and Goldberg, D. E. (2012). *Plasmodium falciparum* heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers. *Nat. Commun.* 3:1310. doi: 10.1038/ncomms2306
- Nelson, R., Sawaya, M. R., Balbirnie, M., Madsen, A. O., Riekel, C., Grothe, R., et al. (2005). Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435, 773–778. doi: 10.1038/nature03680
- Orr, H. T., and Zoghbi, H. Y. (2007). Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* 30, 575–621. doi: 10.1146/annurev.neuro.29.051605.113042
- Painter, H. J., Campbell, T. L., and Llinas, M. (2011). The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. *Mol. Biochem. Parasitol.* 176, 1–7. doi: 10.1016/j.molbiopara.2010.11.014
- Pallarès, I., Iglesias, V., and Ventura, S. (2016). The Rho termination factor of *Clostridium botulinum* contains a prion-like domain with a highly amyloidogenic core. *Front. Microbiol.* 6:1516. doi: 10.3389/fmicb.2015.01516
- Pallares, I., and Ventura, S. (2017). The transcription terminator Rho: a first bacterial prion. *Trends Microbiol.* 25, 434–437. doi: 10.1016/j.tim.2017.03.008
- Pease, B. N., Huttlin, E. L., Jedrychowski, M. P., Talevich, E., Harmon, J., Dillman, T., et al. (2013). Global analysis of protein expression and phosphorylation of three stages of *Plasmodium falciparum* intraerythrocytic development. *J. Proteome Res.* 12, 4028–4045. doi: 10.1021/pr400394g
- Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., et al. (2005). FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21, 3435–3438. doi: 10.1093/bioinformatics/bti537
- Przyborski, J. M., Diehl, M., and Blatch, G. L. (2015). Plasmodial HSP70s are functionally adapted to the malaria parasite life cycle. *Front. Mol. Biosci.* 2:34. doi: 10.3389/fmolb.2015.00034
- Reddy, B. P., Shrestha, S., Hart, K. J., Liang, X., Kemirembe, K., Cui, L., et al. (2015). A bioinformatic survey of RNA-binding proteins in *Plasmodium*. *BMC Genomics* 16:890. doi: 10.1186/s12864-015-2092-1
- Roberg, K. J., Crotwell, M., Espenshade, P., Gimeno, R., and Kaiser, C. A. (1999). LST1 is a SEC24 homologue used for selective export of the plasma membrane ATPase from the endoplasmic reticulum. *J. Cell Biol.* 145, 659–672. doi: 10.1083/jcb.145.4.659
- Sabate, R., Rousseau, F., Schymkowitz, J., Batlle, C., and Ventura, S. (2015a). Amyloids or prions? That is the question. *Prion* 9, 200–206. doi: 10.1080/19336896.2015.1053685
- Sabate, R., Rousseau, F., Schymkowitz, J., and Ventura, S. (2015b). What makes a protein sequence a prion? *PLoS Comput. Biol.* 11:e1004013. doi: 10.1371/journal.pcbi.1004013
- Sant'Anna, R., Fernandez, M. R., Batlle, C., Navarro, S., de Groot, N. S., Serpell, L., et al. (2016). Characterization of amyloid cores in prion domains. *Sci. Rep.* 6:34274. doi: 10.1038/srep34274
- Si, K. (2015). Prions: what are they good for? *Annu. Rev. Cell Dev. Biol.* 31, 149–169. doi: 10.1146/annurev-cellbio-100913-013409
- Singh, G. P., Chandra, B. R., Bhattacharya, A., Akhouri, R. R., Singh, S. K., and Sharma, A. (2004). Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 137, 307–319. doi: 10.1016/j.molbiopara.2004.05.016
- Siomi, H., Matunis, M. J., Michael, W. M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.* 21, 1193–1198. doi: 10.1093/nar/21.5.1193
- Tariq, M., Wegrzyn, R., Anwar, S., Bukau, B., and Paro, R. (2013). Drosophila GAGA factor polyglutamine domains exhibit prion-like behavior. *BMC Genomics* 14:374. doi: 10.1186/1471-2164-14-374
- Tartaglia, G. G., Cavalli, A., Pellarin, R., and Caflisch, A. (2005). Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. *Protein Sci.* 14, 2723–2734. doi: 10.1110/ps.051471205
- Tartaglia, G. G., and Vendruscolo, M. (2008). The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev.* 37, 1395–1401. doi: 10.1039/b706784b
- Tompa, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33, 2–8. doi: 10.1016/j.tibs.2007.10.003
- Toombs, J. A., McCarty, B. R., and Ross, E. D. (2010). Compositional determinants of prion formation in yeast. *Mol. Cell. Biol.* 30, 319–332. doi: 10.1128/MCB.01140-09
- Toombs, J. A., Petri, M., Paul, K. R., Kan, G. Y., Ben-Hur, A., and Ross, E. D. (2012). De novo design of synthetic prion domains. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6519–6524. doi: 10.1073/pnas.1119366109
- Treusch, S., and Lindquist, S. (2012). An intrinsically disordered yeast prion arrests the cell cycle by sequestering a spindle pole body component. *J. Cell Biol.* 197, 369–379. doi: 10.1083/jcb.201108146
- UniProt Consortium (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989
- Williams, A. J., and Paulson, H. L. (2008). Polyglutamine neurodegeneration: protein misfolding revisited. *Trends Neurosci.* 31, 521–528. doi: 10.1016/j.tins.2008.07.004
- Wu, Y., Nelson, M. M., Quail, A., Xia, D., Wastling, J. M., and Craig, A. (2009). Identification of phosphorylated proteins in erythrocytes infected by the human malaria parasite *Plasmodium falciparum*. *Malar. J.* 8:105. doi: 10.1186/1475-2875-8-105
- Yuan, A. H., and Hochschild, A. (2017). A bacterial global regulator forms a prion. *Science* 355, 198–201. doi: 10.1126/science.aai7776
- Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., and Kaneko, I. (2010). Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol. Microbiol.* 75, 854–863. doi: 10.1111/j.1365-2958.2009.07005.x
- Zambrano, R., Conchillo-Sole, O., Iglesias, V., Illa, R., Rousseaus, F., Schymkowitz, J., et al. (2015). PrionW: a server to identify proteins containing glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Res.* 43, W331–W337. doi: 10.1093/nar/gkv490
- Zhang, M., Mishra, S., Sakthivel, R., Rojas, M., Ranjan, R., Sullivan, A. J. Jr., et al. (2012). PK4, a eukaryotic initiation factor 2alpha(eIF2alpha) kinase, is essential for the development of the erythrocytic cycle of *Plasmodium*. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3956–3961. doi: 10.1073/pnas.1121567109

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pallarès, de Groot, Iglesias, Sant'Anna, Biosca, Fernández-Busquets and Ventura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.